Property Taxes and Home Prices: A Tale of Two Cities

ChongEn Bai
Tsinghua SEM

Qi Li*
Taxas A&M

Min Ouyang
Tsinghua SEM

September 2012

Abstract

We explore the influence of property taxes on home prices, taking advantage of a policy experiment of property taxes in Shanghai and in Chongqing starting from January 2011. An econometric model is proposed to estimate hypothetical home prices in the absence of property taxes for Shanghai and Chongqing using home prices in other cities and provinces. The model suggests the OLS generates consistent estimators when the price series are non-stationary. We apply the model to a panel of average home prices of 31 cities and provinces in China, and find the property-tax experiment lowered the Shanghai average home price by 15% but raised the Chongqing one by 11%. An examination of the policy details and data on prices by home types suggests the post-treatment price increase in Chongqing can be driven by a spillover effect from high-end to low-end properties.

JEL codes: C10, H73, H22.

Key words: property tax, home price, unit-root process, OLS estimators, spillover effect.

*We thank participants at the 14th NBER-CCER conference on China and the World Economy, especially James Porteba, for useful comments. The China Data Center at Tsinghua University provides the data. Errors are ours. Please do not cite or circulate without permission.

1 Introduction

How do property taxes influence home prices? The literature on local public finance says the effect should be strictly negative, as long as property taxes are, at least partially, capitalized.[1] Intuitively, property taxation imposes additional user costs on a property and thus reduces its value. Under full capitalization, differences in home prices exactly equal the present discounted value of variations in expected property tax liabilities. To see this, suppose a property has a finite life span of n years. Let P to be its market value. Y_t is the inflow of property value in year t. i is the interest rate and τ the property tax rate. Under standard assumptions,

$$P = \sum_{t=1}^{n} \frac{(Y_t - \tau P)}{(1+i)^{t-1}}. \tag{1}$$

Apparently, P declines in τ.

However, testing the influence of property taxes on home prices involves several difficulties. Firstly, the causality can run from P to τ. If the local government targets a fixed amount of tax revenue, then lower tax rates can be imposed on communities with higher home values. Secondly, Y_t, i, and other factors are hard to control for. For example, Y_t is associated with the quality of local public services, monetary policies, inflation, and public expectations (Porteba (1984)). All these factors are hard to fully identify. The literature has pointed out that, when property taxes are used to finance local public services like in the U.S., higher tax rate is associated with higher P by improving the quality of public goods (Rosen and Fullerton (1977)). To avoid biases arising from these endogeneity problems, some authors use natural experiments derived from exogenous policy changes (for example, Rosen (1982)). Nonetheless, even if changes in τ are exogenous, it remains challenging to fully control for Y_t, for i, and for other factors.

This paper estimates the influence of property taxes on home prices, taking advantage of a property-tax experiment implemented in China at the end of January 2011, in two cities only – Shanghai and Chongqing. Unlike many other countries, there has been no property taxes in China until then. Thus, in addition to having an exogenous change in τ, our study offers several advantages. Firstly, since property taxes have not been a major source of Chinese governments' tax revenue and are not used to finance local public goods, it avoids a standard bias in this literature that higher property taxes are associated with better public goods.[2] Secondly, Shanghai and Chongqing, one located on the east

[1] This was first formally developed and tested by Oates (1969). Many authors followed including Rosen and Fullerton (1977), Rosen (1982), Palmon, and Smith (1998), and Feldman (2010).

[2] Both Shanghai and Chongqing governments use the proper-tax revenue to finance the construction of subsidized rental houses for the poor. Since these houses are still under construction by the end of our sample period, this should

coast and the other on the south west of China, differ in their political and economic characteristics; comparison of the results for the two cities helps to address the sample selection bias – also a standard one in applying policy experiments. Thirdly and most importantly, we can use home prices in other cities/provinces to control for potential changes in Y_t, in i, and in other factors for Shanghai and Chongqing, instead of identifying variations in each factor. In particular, we estimate hypothetical home prices in the absence of property taxes in the treatment group using home prices in the control group, compare hypothetical prices with actual prices, to identify the treatment effect of the property-tax experiment.

This approach, motivated by Hsiao, Ching, and Wand (hereafter HCW)(2011), is similar in spirit to the conventional difference-in-difference (DID) approach. Yet, it allows for more flexibilities in the estimation. To see this, suppose Y_t changes before and after the policy experiment. The DID approach assumes the treatment and control groups share exactly the same change in Y_t as well as bear the same influence, so that taking differences has it removed. These can hardly apply to local home-price variations in China. Suppose that, an expansionary fiscal policy drives up home prices in all cities like the 2008 China Fiscal Stimulus Plan. It is possible that home prices rise by more in Shanghai than in Jiangsu or vice versa because, in China, local governments' economic powers vary so that their responses to macro policies also vary. Failure to incorporate such regional heterogeneity can falsely attribute home-price changes driven by other factors to the property-tax experiment, creating biases on the estimators.

Instead, our approach focuses on estimating the *correlation* between home prices in the treatment group and those in the control group. Hence, it allows for the impact of underlying factors to vary by city/province. Also, our approach puts more weight on control cities/provinces more relevant to the treatment cities, unlike the DID approach that assigns the same weight to each control-group member. For example, Jiangsu, as a neighborhood province of Shanghai, gets more weight than Heilongjiang when both serving as control provinces for Shanghai. These details are carefully presented in an econometric model in Section 2. The model extends from HCW (2011) without relying on a key assumption. We show that, as long as the price series are non-stationary, the OLS estimation generates consistent estimators for the correlation, for hypothetical prices, and therefore for the treatment effect of the property-tax experiment.

We apply this approach to data. Perhaps surprisingly, we find totally opposite effects of property taxation on home prices in Shanghai and in Chongqing. The estimates suggest the property-tax experiment has *lowered* the Shanghai average home price by 15% but *raised* the Chongqing one by

not influence the value of commercial housing and thus cannot bias our estimates.

11%. These results stay quite robust to various estimation specifications and to stationary versus non-stationary data. A close examination of the policy shows taxation specifics differ for the two cities. In Chongqing property taxes are mainly imposed on high-end properties including single family houses, big apartments, and those much more expensive than the city average. We propose the positive effect of property taxes on home prices in Chongqing, opposite to that in Shanghai and counter-intuitive according to the literature of property-tax capitalization, is driven by a spillover effect from high-end to low-end properties. Intuitively, people quit buying high-end homes, turn to low-end ones to avoid future property-tax payments. This lowers prices of high-end houses but raises those of low-end ones. A simple examination of data on prices by home type supports our hypothesis.

The paper proposes an estimation approach most applicable to policy experiments and to non-stationary data, which should be a valuable tool for studying Macroeconomic policies. Moreover, it provides an important suggestion for housing policies currently under intensive discussion in China. In the past ten years China has experienced a dramatic increase in home prices. The magnitude has been astonishing: it is said that the national average home price has tripled from 2005 to 2009. The increase in home prices has dominated that in the household income: the ratio of median housing price to median annual disposable household income, a standard measure for housing affordability, equals 27 in Beijing, five times of the international average.[3] Under such circumstances, this policy experiment was implemented at the purpose of exploring property taxation as a policy tool to lower home prices. Although this paper cannot evaluate many other impacts of property taxation on, for example, local public services, national investment rate, and social welfare, it does offer an important piece of advice for future property-tax policy. That is, property taxation should be implemented very carefully if it is for the purpose of stabilizing home prices. In particular, we should be cautious in following Chongqing by imposing discriminative property taxes based on home types, because this can generate a spillover effect and cause consequences opposite to what the government intends for.

The rest of the paper is organized as follows. Section 2 lays out the econometric model. Section 3 describes the data. The estimation results are discussed in Section 4. Section 5 explores the potential spillover effect in Chongqing. We conclude in Section 6.

[3]See the 8th annual demographia international housing affordability survey published by the Wendell Cox Consultancy. A ratio below 3.0 is considered as "affordable" and that above 5.1 is "severely unaffordable". This ratio ranges from 2.7 to 3.1 for the U.S..

2 The Model

Let P_{it}^1 and P_{it}^0 denote city i's (average) home price in period t with and without property taxes, respectively. The property tax policy intervention effect to city i at time t is

$$\Delta_{it} = P_{it}^1 - P_{it}^0. \tag{2}$$

However, often we do not simultaneously observe P_{it}^0 and P_{it}^1. The observed data are in the form

$$P_{it} = d_{it} P_{it}^1 + (1 - d_{it}) P_{it}^0, \tag{3}$$

where $d_{it} = 1$ if the city i has the property tax (under treatment) at time t, and $d_{it} = 0$ otherwise.

Following HCW (2011) we assume that there exists a $K \times 1$ vector of *unobservable* common factors f_t that drives home prices of all cities to change over time. In our application, these can be national economic growth, macro policies, borrowing opportunities, environmental improvements, and changes in public expectations. Apparently, in this case f_t is more likely to be non-stationary, its components can either be a unit root process without drift, or a unit root process with drift. We consider the case that all cities did not have property taxes for $t = 1, ..., T_1$, i.e., $t \leq T_1$ corresponds to pre-treatment period.

$$P_{it}^0 = \alpha_i + b_i' f_t + u_{it}, \qquad i = 1, ..., N, \ t = 1, ..., T_1, \tag{4}$$

where α_i is an individual specific intercept, b_i is a factor loading vector of dimension $K \times 1$, u_{it} is a stationary I(0) error term. Note b_i can differ by i. We consider the usual case that $K < N$ so that a few common factors affect home prices of different cities.

Starting from period $T_1 + 1$, government implements a property-tax experiment only to one city. Without loss of generality, we assume this is the first city,

$$P_{1t}^1 = \alpha_1 + b_1' f_t + \Delta_{1t} + u_{1t}, \qquad t = T_1 + 1, ..., T, \tag{5}$$

where Δ_{1t} is the treatment effects which captures the impact of property tax on home price in city 1 after the implementation of the new property tax.

Because for $t \geq T_1 + 1$, the new property tax is implemented only to city 1, all other cities still do not have property taxes. Therefore, for all time horizon, we have

$$P_{it}^0 = \alpha_i + b_i' f_t + u_{it}, \qquad i = 2, ..., N \text{ and } t = 1, ..., T. \tag{6}$$

Let $P_t = (P_{1t}, ..., P_{Nt})'$ be an $N \times 1$ vector of P_{it}s at time t. Since there is no property tax intervention before T_1, then for $t \leq T_1$ the observed P_t takes the form

$$P_t = P_t^0 = \alpha + Bf_t + u_t \qquad \text{for } t = 1, ..., T_1, \tag{7}$$

where $\alpha = (\alpha_1, ..., \alpha_N)'$, B is a $N \times K$ matrix with the i^{th} row given by b_i', and $u_t = (u_1, ..., u_N)'$.

Since at time $T_1 + 1$, there is a new property tax imposed to city 1, hence, from time $T_1 + 1$ on we have

$$P_{1t} = P_{1t}^1 \qquad \text{for } t = T_1 + 1, ..., T. \tag{8}$$

We assume that other cities are not affected by the property tax implementation at the first city. Therefore, for all time horizon

$$P_{it} = P_{it}^0 = \alpha_i + b_i' f_t + u_{it}, \qquad \text{for } i = 2, ..., N \text{ and } t = 1, ..., T. \tag{9}$$

Since P_{1t}^0 is not observable for $t \geq T_1 + 1$, we need to estimate the counterfactual home price P_{1t}^0 when $t > T_1$. If T and N are large, method of Bai and Ng (2002) can be used to identify the number of common factor, K, and estimate f_t along with B by the maximum likelihood approach. HCW (2011) suggest an alternative method by using $\tilde{P}_t = (\tilde{P}_{2t}, ..., \tilde{P}_{Nt})'$ in lieu of f_t to predict P_{1t}^0 for post-treatment period. The validity of HCW's approach depends on a condition that: there exists a $N \times 1$ vector $a = (1, -\gamma')' = (1, -\gamma_2, ..., -\gamma_N)'$ such that $a'B = 0$ (i.e., $a \in \mathcal{N}(B)$, the null space of B) and that one can consistently estimate a.[4] This condition is stated as an assumption in HCW (their assumption 6). HCW (2011) consider the stationary data case, their assumption 6 is equivalent to assuming that the conditional mean function of u_t, conditional on $\tilde{P}_t$, is a linear function of $\tilde{P}_t$. In our case because f_t and P_t are I(1) variables, we do not need to make any linear conditional mean functional form assumption, we will prove a result similar to assumption 6 of HCW in Proposition 2.1 below, the proof of Proposition 2.1 is given at the Appendix A of our paper.

Proposition 2.1 *Under assumptions A1 to A4 given below, we have*

(i) For fixed K and N ($K < N$), there exists a unique solution $(\gamma_1, \gamma_2, ..., \gamma_N) \equiv (\gamma_1, \gamma')$ that minimizes

$$E\left[(P_{1t}^0 - \gamma_1 - \gamma' \tilde{P}_t)^2\right],$$

and that $a = (1, -\gamma') \in \mathcal{N}(B)$, i.e., $a'B = 0$.

(ii) The OLS estimator $\hat{\gamma}_1$ and $\hat{\gamma}$ based on $P_{1t} = \gamma_1 + \gamma' \tilde{P}_t + \epsilon_{1t}$, $t = 1, ..., T_1$, are consistent

[4] In the application of the difference-in-difference approach, it is equivalent to have a B with identical rows so that $a = (1, -\gamma')' = (1, -\frac{1}{N-1}, ..., -\frac{1}{N-1})'$ satisfies $a'B = 0$ and removes the impact of the latent factors. Our approach offers more flexibility by allowing rows in B to differ, so that the existence of a must be addressed.

estimators of γ_1 and γ.

Proposition 2.1 states that one can consistently estimate $(\gamma_1, \gamma')'$ by the least squared method of regressing P^0_{1t} on $(1, \tilde{P}_t)$ using the pre-treatment data.

Recall that $P_t = (P_{1t}, \tilde{P}'_t)'$, where $\tilde{P}_t = (P_{2t}, ..., P_{Nt})'$, and P^0_t is generated by

$$P^0_t = \alpha + Bf_t + u_t$$

for $i = 1, ..., N$ and $t = 1, ..., T_1$. If $a = (1, -\gamma')'$ satisfies $a'B = 0$. Then we have that

$$a'P_t \equiv P_{1t} - \gamma'\tilde{P}_t = a'\alpha + a'u_t$$

because $a'B = 0$ so that the common factors are dropped out from the right-hand-side of the above equation. Rearranging terms we obtain

$$P_{1t} = \gamma_1 + \gamma'\tilde{P}_t + \epsilon_{1t}, \tag{10}$$

where $\gamma_1 = a'\alpha$ and $\epsilon_{1t} = a'u_t = u_{1t} - \gamma'\tilde{u}_t$ with $\tilde{u}_t = (u_{2t}, ..., u_{Nt})'$.

Because ϵ_{1t} is I(0) and $\tilde{P}_t$ is I(1), equation (10) implies that P_{1t} and $\tilde{P}_t$ are cointegrated. However, since some or all the components P_{jt} contains drift terms, for $j = 2, ..., N$, some of the P_{jt}s are asymptotically collinear. One way to avoid the near multicollinear problem is to add a time trend variable to the right-hand-side the regression model, after adding a time trend regressor, one can also de-trend the right-hand-side P_{jt}, for $j = 2, ..., N$, see Xiao (2001). Or simply add a time trend variable to the right-hand-side of the regression equation (20). Which method to adopt is unimportant to our purpose since our interest is not on the accurate estimation of the individual coefficient γ_j, $j = 2, ..., N$, but rather we are interested in estimating prediction component $\gamma_1 + \gamma'\tilde{P}_t$. Let $\hat{\gamma}_1$ and $\hat{\gamma}$ be the OLS estimates of γ_1 and γ, respectively. Then $\hat{g}(\tilde{P}_t) \equiv \hat{\gamma}_1 + \hat{\gamma}'\tilde{P}_t$ consistently estimates $g(\tilde{P}_t) \equiv \gamma_1 + \gamma'\tilde{P}_t$ in the sense that $T_1^{-1}\sum_{t=1}^{T_1} \hat{g}(\tilde{P}_t) - T_1^{-1}\sum_{t=1}^{T_1} g(\tilde{P}_t) = O_p(T_1^{-1/2}) = o_p(1)$. So that we can consistently estimate the average treatment effects based on (20). In practice if multicollinearity is severe, one can simply remove some of components of $\tilde{P}_t$ to reduce multicollinearity. Indeed for the empirical data we have, we will show usually N equals to 3 or 4 are sufficient to generate excellent fit based on (10).

Note that the new error ϵ_{1t} is in general correlated with $\tilde{P}_t$ because ϵ_t depends on all the original error $(u_{1t}, ..., u_{Nt})'$. Even so we show in Appendix A that the least squares estimators based on (10), say $(\hat{\gamma}_1, \hat{\gamma}')'$ has the property that $\hat{\gamma} \xrightarrow{p} \gamma$ and that $a'B = (1, -\gamma')B = 0$. The reason is that the regressor $\tilde{P}_t$ is an I(1) process, and the error ϵ_{1t} is an I(0) stationary process. It is well known that in such cointegrated models, mild endogenous regressors do not affect the consistency result of the least squares estimators.

Since the new property tax is imposed only to city 1 so that $P_{1t} = P_{1t}^1$ for $t = T_1 + 1, ..., T$, but this new tax on city 1 does not affect $\tilde{P}_t$. Hence, we have $\tilde{P}_t = \tilde{P}_t^0$ for all $t = 1, ..., T$. We will drop the superscript 0 in $\tilde{P}_t^0$ because $\tilde{P}_t = \tilde{P}_t^0$ for all $t = 1, ..., T$. We estimate $(\gamma_1, \gamma')'$ based on the following linear regression model

$$P_{1t}^0 = \gamma_1 + \tilde{P}_t'\gamma + \epsilon_{1t}, \qquad t = 1, ..., T_1, \tag{11}$$

where γ_1 is a scalar parameter, and $\gamma = (\gamma_2, ..., \gamma_N)'$ is a $(N-1) \times 1$ vector of parameter, ϵ_{1t} is a stationary $I(0)$ error term.

Note that in the absent of property tax, we would have

$$P_{1t}^0 = \gamma_1 + \tilde{P}_t'\gamma + \epsilon_{1t}, \qquad t = T_1 + 1, ..., T. \tag{12}$$

However, since there is a property tax imposed to city 1, P_{1t}^0 is not observable for $t > T_1$. Nevertheless, because there is no property tax on $\tilde{P}_t$ for all t, (12) suggests that we can estimate P_{1t}^0 by $\gamma_1 + \tilde{P}_t'\gamma$, provided that we can consistently estimate the unknown parameters γ_1 and γ. We will follow this approach as suggested by HCW. In order to establish the consistency result presented in proposition 2.1, below we first make some assumptions.

Assumption 1. (i) The factor follows unit root process $f_t = c + f_{t-1} + v_t$, where c is a $K \times 1$ vector of constants, v_t is zero mean I(0) process. (ii) $\lim_{T\to\infty} T^{-1} \sum_{t=1}^T \sum_{s=1}^T E(v_t v_s') = \Omega$, where Ω is a $K \times K$ positive definite matrix.

Assumption 2. (i) $P_t = \alpha + Bf_t + u_t$, Write $B = \begin{pmatrix} b_1' \\ \tilde{B} \end{pmatrix}$, where b_1' is the first row of B and $\tilde{B}$ is of dimension $(N-1) \times K$. There exists a squared sub-matrix of $\tilde{B}$, say B_K, such that B_K is invertible. (ii) u_t is zero mean I(0) process satisfying $Var(u_t|\tilde{P}_t) = \sigma^2(\tilde{P}_t)$ with $\sigma^2(\tilde{P}_t)$ is bounded from both below and above by positive constants.

Assumption 3. We re-index the time index by $j = t - T_1$ so that $j = 1, ...T_2 = T - T_1$ when $t = T_1 + 1, ..., T$, for the de-trended process $\tilde{P}_j^* = \tilde{P}_{j-1}^* + \eta_j$, where η_j is a weakly dependent (mixing) stationary process. Define $B_{T_2}(r) = T_2^{-1/2} \sum_{j=1}^{[T_2 r]} \eta_j$, where $r \in [0, 1]$, $[a]$ denotes the integer part of a. Then the following functional central limit theorem holds for the $B_{T_2}(r)$ process: $B_{T_2} \Rightarrow W$, where $W(\cdot)$ is a $(N-1) \times 1$ vector of Brownian motion with zero mean and covariance matrix given by $\Sigma = \lim_{T_2\to\infty} Var(T_2^{-1/2} \sum_{j=1}^{T_2} \eta_j)$, here $\Rightarrow$ denotes weak convergence.

Assumption 4. Both T_1 and T_2 are large, where $T_2 = T - T_1$, and that $T_2/T_1 = O(1)$ as $T_1 \to \infty$,

$T_2 \to \infty$.

Assumption 1 states that the common factors follow unit root processes with drifts. Further, the positive definite of Ω implies that different components of f_t are not cointegrated. Assumption 2, together with assumption 1, imply that P_t also follows unit root processes. Assumption 2 assumes that the error u_t is an I(0) process, while P_t is an I(1) non-stationary process. Assumption 3 is satisfied by many weakly dependent mixing processes, the conditions for the multivariate functional central limit theorem for partial sums of weakly dependent random vectors can be found in de Jong and Davidson (2000). Assumption 4 allows for either T_1 is much larger than T_2, or T_1 and T_2 have the same order of magnitude. We do not deal with the case that T_2 is much larger than T_1 as it is unlikely that one wants to evaluate a policy effects with a policy implemented in the long past, and with relative small data available for pre-treat period.

Recall that the property tax (treatment) effects at time t is defined as

$$\Delta_{1t} = P^1_{1t} - P^0_{1t}, \qquad\qquad t = T_1 + 1, ..., T. \tag{13}$$

When $t \geq T_1 + 1$, we only observe P^1_{1t}, not P^0_{1t}. Hence, Δ_{1t} defined in (13) is not observable. However, using (12) and (13) we can predict P^0_{1t} by $\gamma_1 + \tilde{P}_t'\gamma$ for $t = T_1 + 1, ..., T$. In practice γ_1 and γ are unknown, these unknown parameters can be consistently estimated by the least squares method. Hence, we can predict P^0_{1t} by $\hat{P}^0_{1t} = \hat{\gamma}_1 + \tilde{P}_t'\hat{\gamma}$, where $\hat{\gamma}_1$ and $\hat{\gamma}$ are the least squares estimates of γ_1 and γ based on the pre-treatment period data. Therefore, we can estimate the treatment effects by

$$\hat{\Delta}_{1t} = P^1_{1t} - \hat{P}^0_{1t}, \qquad\qquad t = T_1 + 1, ..., T, \tag{14}$$

and we estimate the average treatment effect by

$$\hat{\Delta}_1 = \frac{1}{T_2} \sum_{t=T_1+1}^{T} \hat{\Delta}_{1t}, \tag{15}$$

where $T_2 = T - T_1$.

The asymptotic properties of $\hat{\Delta}_{1t}$ and $\hat{\Delta}_1$ are given in Propositions 2.2 and Proposition 2.3 below.

Proposition 2.2 *Under Assumptions 1 to 4,* $\hat{\Delta}_{1t} = \Delta_{1t} + \epsilon_{1t} + O_p(T_1^{-1/2})$ *for any* $t \in \{T_1 + 1, ..., T\}$.

Note that in Proposition 2.2 we do not require that the treatment effect Δ_{1t} to be a stationary process. However, if Δ_{1t} is a stationary process, define $\Delta_1 = E(\Delta_{1t})$, then one can show that Δ_1 can be consistently estimated by $\hat{\Delta}_1$. We will make an additional regularity condition before establishing the consistency of $\hat{\Delta}_1$ for Δ_1.

Assumption 5. Δ_{1t} is a weakly dependent stationary (ergodic) process such that $T_2^{-1}\sum_{t=T_1+1}^{T}\Delta_{1t} \xrightarrow{p} \Delta_1$ as $T_2 = T - T_1 \to \infty$.

Proposition 2.3 *Under Assumptions 1 to 5, we have*

$$\hat{\Delta}_1 \xrightarrow{p} \Delta_1,$$

where $\hat{\Delta}_1$ *is defined in (15).*

Up to know we have only considered the non-stationary pricing data case. If one considers the annual growth rate of home price, say $p_t = \ln P_t - \ln P_{t-12}$ (home price annual growth rate), then it is possible that p_t is a stationary variable. Using panel data to estimate treatment effects with stationary variables was considered by Hsiao, Ching and Wand (2011). The estimation method is the same as above, the consistency of the estimation results requires an additional assumption that $E(\epsilon_{it}|\tilde{y}_t)$ is a linear function in $\tilde{y}_t$, see assumption 6 of HCW. In the empirical applications reported in the next section, as a robust check, we also use home price growth rate, as well as de-trended data, to analyze the property tax effects on home price.

3 The Data

To apply the model to data, we need a panel of home prices across cities/provinces including our treatment cities, ideally at a high frequency to obtain sufficient sample size. The National Development and Reform Commission (NDRC) of China provides such data. It reports monthly average per-square-meter prices of all existing commercial residential properties across 27 provinces and four municipalities nationwide. This dataset is compiled from information on all housing transactions reported to the government.

An alternative data is from the National Bureau of Statistics (NBS) that publishes monthly home-price indexes for 70 cities. We prefer the NDRC data because it provides a much bigger sample size by starting from March 1998. More importantly, the NBS series before and after January 2011 – exactly the month when the policy experiment was carried out – are not directly comparable due to a significant change in the data compiling process. This suggests the NBS series are not appropriate for our approach of using pre-treatment prices to estimate post-treatment prices in the absence of the treatment effect. However, in later section we will examine the post-treatment series of the NBS data for a further exploration of our findings.[5]

[5] Some NBS series begin with March 2009 while others start from January 2011. It is based on a nation-wide sample of 10,000 newly constructed and secondary properties compiled from reports by real estate developers.

Hence, we examine the NDRC series as a panel of home price across 31 cities/provinces, including Shanghai and Chongqing, from March 1998 to March 2012. Proposition 2.1 states that, when P_t's are I(1) variables, the OLS approach consistently estimates the average of P_{1t}^0 using $\tilde{P}_t$ without relying on a key assumption of HCW(2011). Accordingly, as the first step we test for the existence of unit roots in the NDRC home price series. Table 1 reports the Mackinnon approximated p-values from the augmented Dickey-Fuller tests on 31 price series. Here the log-first difference is taken as that between the present home price in log levels and that 12 months ago, namely, the annual price growth. The last three columns of Table 1 show that, in log first differences, we *can* reject the unit-root hypothesis with or without a drift for almost all series. In log levels, however, we *cannot* reject the unit-root hypothesis for 30 out of the 31 provinces/cities when the test specification excludes a drift or a trend; this number is 26 with a drift, and only 11 with a time trend.

We interpret the results in Table 1 as follows: average annual growths in home prices are mostly stationary; average home prices in log levels are mostly I(1) series, although for many provinces/cities, controlling for a time trend renders the series stationary. Based on such findings, we conduct our estimation with average home prices measured in log levels, to satisfy Assumptions 1 and 2 and to obtain consistent estimates as suggested by Proposition 2.1. We also experiment with including a time trend in the regression to check for the robustness of our results.

4 Estimation of the Treatment Effect

We apply the data described in Section 3 to the model specified in Section 2 to evaluate the treatment effect of property taxes. In particular, we use the home-price data from March 1998 to January 2011 to estimate:

$$P_{1t} = \gamma_1 + \gamma' \tilde{P}_t + \gamma_2 D_{2008} + \gamma_3 t + \epsilon_{1t}, \tag{16}$$

P_{1t} is home price in Shanghai or in Chongqing. $\tilde{P}_t$ are home price in control provinces/cities. (16) is similar to (10) except for two additional right-hand-side terms. D_{2008} is a post-2008 dummy that equals one after November 2008 and equals zero otherwise; it is supposed to capture the influence of the 2008 Fiscal Stimulus Plan widely believed to have largely raised home prices nation-wide. D_{2008} improves the fit of the estimation, although taking it off does not change our results qualitatively or quantitatively. t is an *optional* time trend; as suggested in Section 2, we experiment with this time trend to avoid potential multicollinearity arising from drift terms in the I(1) processes.

4.1 A Tale of Two Cities

Equation (16) is estimated by OLS for Shanghai and for Chongqing separately. The sample size of each regression is 155. To obtain sufficient degrees of freedom, we select only a few provinces/cities into each control group by choosing those giving the best fit of the estimation. Table 2 reports the results. The control provinces/cities are Jiangsu, Zhejiang, Heilongjiang, and Sichuan for Shanghai, and Jiangsu, Zhejiang, Beijing, and Sichuan for Chongqing. Adjusted R-squares and F-statistics are very high for both cities, implying home prices in control provinces/cities serve as good predictors for those in the treatment cities. The estimation results of (16) with and without a time trend are presented in Columns 2-4 and 5-7 of Table 2, respectively. In general including a time trend does not show much influence on the point estimates or the standard errors.

4.1.1 The estimated weights for control provinces

The selection of control provinces as listed in Table 2 reflects a geographical clustering effect: for example, Jiangsu and Zhejiang are both neighborhood provinces for Shanghai, and so is Sichuan for Chongqing. People living in the same region can have similar housing purchasing behaviors by sharing the same climate, culture, income profile, and spending habits. Also, it turns out that Beijing home prices have prediction powers for those in Chongqing but not for those in Shanghai. This is possibly because Chongqing is known for its strong political characteristics, so that its economic outcome is also tightly related to that in the capital city Beijing.

Table 2 reports some of the estimated coefficients to be negative. For example, the estimated weight of Sichuan is negative when serving as a control province for Shanghai. While this isn't problematic statistically, one may question its economic meaning. Note that Shanghai and Sichuan are located geographically far from each other – one by the east coast and the other in the Southwestern inland. Therefore, it is possible for home prices of these two regions to move sometimes in opposite directions (after controlling for other factors) due to, for example, migration from inland to the coast or the adoption of policies in favor of high-tech sector that attracts IT companies to move from the coast to the Southwestern inland.[6]

Also note that, for Chongqing, the coefficient of the time trend is estimated as negative and statistically significant. This is because the average home price in Chongqing displayed a clear downward trend before 2002 (See Figure 1). In other words, the Chongqing home prices were declining before

[6]For example, Chengdu, as the capital city of Sichuan province, has developed a high-tech district starting from 1988, providing tax exemptions and subsidies to high-tech companies. In recent years, many IT companies have moved their headquarters from Shanghai to Chengdu.

2002, and started to rise continually only after 2002. This downward trend before 2002 is also present for some of the other provinces/cities, but appears much milder. As a matter of the fact, taking off the data before January 2002 renders this time-trend coefficient for Chongqing positive and statistically significant. Nonetheless, our results remain robust to keeping this trend, taking off this trend, or allowing this trend to differ before and after January 2002.

4.1.2 The estimated treatment effect

Next, we construct $\hat{P}^0_{1t}$, the hypothetical home price in the absence of property taxes, using the estimated weights reported in Table 2, to generate estimates for the treatment effect defined as $\hat{\Delta}_{1t} = P^1_{1t} - \hat{P}^0_{1t}$. The number of observations pre-treatment (T_1) is 155 and that post-treatment (T_2) is 14, satisfying Assumption 4 that T_1 should dominate T_2. Table 3 lists P^1_{1t}, $\hat{P}^0_{1t}$, and $\hat{\Delta}_{1t}$ starting from February 2011. Panels A and B of Table 3, respectively, report those using weights estimated without and with a time trend. Although our theory suggests the true treatment effect can be consistently estimated using the average treatment effect under certain assumptions, in this case we cannot examine the statistical significance of $\hat{\Delta}_{1t}$ as T_2 (14 observations) is too small to evaluate its time-series properties.

Nonetheless, the estimated treatment effects give an interesting tale of two cities. Surprisingly, the property-tax experiment has had *opposite* effect on the average home price in Shanghai and in Chongqing. Table 3 shows that, starting from February 2011, the estimated effect is strictly negative for Shanghai but positive for Chongqing. The average treatment effect for Shanghai is -0.1532 based on weights estimated without a time trend, and -0.1517 based on those with a time trend; by contrast, the corresponding numbers are 0.1151 and 0.1063 for Chongqing. Put intuitively, the property-tax experiment has *lowered* the Shanghai average home price by about 15% but *raised* the Chongqing average home price by about 11%.

The opposite treatment effects for the two cities are most apparent in Figure 1, which plots P^1_{1t} and $\hat{P}^0_{1t}$. Because the estimated coefficient on the time trend is statistically insignificant for Shanghai but significant for Chongqing, the plotted $\hat{P}^0_{1t}$ is constructed based on weights estimated without a time trend for Shanghai but those with a time trend for Chongqing. The top two panels show that, before the treatment, $\hat{P}^0_{1t}$ closely tracks P^1_{1t} for both cities while, after the treatment, $\hat{P}^0_{1t}$ goes above P^1_{1t} for Shanghai but below P^1_{1t} for Chongqing. The sizable post-treatment gaps between P^1_{1t} and $\hat{P}^0_{1t}$ are most clear in the bottom two panels that present the post-treatment series only. According to Figure 1, the average home price would have been higher in Shanghai but lower in Chongqing in the absence of property taxes.

4.2 Robustness and Consistency Check

Before exploring explanations for our results, we conduct two robustness checks. The first is to check the prediction power of the control group. One cannot say the prediction is accurate just because $\hat{P}_{1t}^0$ tracks P_{1t}^1 closely before the treatment, because $\hat{P}_{1t}^0$ is itself generated using weights estimated by regressing P_{1t}^1 on $\tilde{P}_t$. To assess the prediction accuracy, we cut the data by 14 months, and repeat the OLS estimation of (16) using the truncated data. More specifically, we re-estimate (16) using data from March 1998 to November 2009 to form predictions for home prices in Shanghai and in Chongqing starting from December 2009. In this case, the difference between P_{1t}^1 and $\hat{P}_{1t}^0$ from December 2009 to January 2011 reflects the prediction accuracy. Figure 2 plots P_{1t}^1 and $\hat{P}_{1t}^0$ from December 2009 to March 2012 estimated based on the truncated data: P_{1t}^1 and $\hat{P}_{1t}^0$ still closely track each other for the first 14 months, implying prediction accuracy; P_{1t}^1 goes below $\hat{P}_{1t}^0$ in Shanghai but above $\hat{P}_{1t}^0$ in Chongqing starting from February 2011, suggesting opposite treatment effects.

Secondly, we follow HCW(2011) to estimate the treatment effects using stationary series. So far, all estimations are conducted in log levels that are I(1) series, which, according to Proposition 2.1, gives consistent estimators when applied to the OLS estimation. However, HCW (2011) *assumes* consistent estimators exist with stationary series. Accordingly we repeat the OLS estimation of (16) in log-first differences (growths). Here, growth is measured as annual growth, the change in log levels compared with the corresponding month in the previous year. Table 3 reports the results. Adjusted R-square's are all above 0.80, implying good fits. Controlling for a time trend does not change the estimated coefficients by much.

The estimated treatment effects in annual growths are reported in Table 5. Corresponding P_{1t}^1 and $\hat{P}_{1t}^0$ are plotted in Figure 3. Apprently, the log-level estimation results carry over to the annual-growth estimation. In Table 5, the treatment effect is negative for Shanghai but positive for Chongqing. In Figure 3, the line of hypothetical prices move above that of actual prices for Shanghai but below that for Chongqing.

To check whether the estimated treatment effects in annual growth are consistent with those in log levels, we convert the estimated log-level prices into annual growths by taking log-first differences between the present month and the corresponding month in the previous year. Figure 4 plots actual price growths (solid lines), hypothetical price growths based on the annual-growth estimation (long dashed lines), and hypothetical price growths converted from hypothetical log-level prices based on the log-level estimation (short-dashed lines). Apparently, the long-dashed line and the short-dashed line follow each other for both cities, suggesting the treatment effect is negative for Shanghai but positive for Chongqing.

5 The Price Spillover

Why has the property-tax experiment generated totally opposite effects on home prices for the two cities? A close look at the tax implementation details of the two cities shows that their policy specifics differ greatly, especially in the following three aspects.

Firstly, not all houses are taxable: in Shanghai the property taxes are imposed on *all* newly purchased houses except for local residents' first purchases; but, in Chongqing, they are imposed on high-end houses only – namely – single family houses (either newly purchased or not) and those newly purchased at a price at least twice as much as the city average. Secondly, not the entire values of the taxable houses are actually taxed: in Shanghai each family member can claim a tax exemption on values associated with 60 square meters of the taxable house; in Chongqing, values associated with the first 180 square meters are tax-exempt for single family houses and, for other taxable houses, it is 100 square meters. Lastly, the property-tax rate is different: in particular, in Shanghai it is 0.6% for taxable houses at least twice as expensive as the city average and 0.4% for others; in Chongqing, the tax rate is 1.2% for taxable houses at least four times as expensive as the city average, 1% for those at least three times as expensive as the city average, and 0.5% for others; moreover, the tax base is taken as 70% of the taxable value in Shanghai but 100% in Chongqing.

Which of the above three aspects helps to explain our results? Lower tax rate in Shanghai cannot help much as, theoretically speaking, home prices cannot possibly fall under lower taxes but rise under higher taxes. The difference in the tax object and in the exemption value can be potential causes. With property taxes imposed on high-end houses only, it is possible for home buyers to turn to lower-end properties instead; similarly, people might start buying houses smaller than the tax-exempt size at the purpose of avoiding property-tax payments. This would raise the price of smaller or lower-end houses but lowers that of bigger or higher-end ones. If the price increase of lower-end (or smaller) houses dominates the price decline of higher-end (or bigger) ones, then the city average home price might rise instead, just as what we have identified for Chongqing. In short, we propose that the post-treatment home-price increase in Chongqing comes from a compositional change in prices of various property types driven by a spillover effect from higher end to lower end.

To explore this hypothesis, we turn to the NBS data as the NDRC panel does not provide information on prices by home type. Starting from January 2011, the NBS publishes city-level price indexes for three categories of houses: those smaller than 90 square meters, of sizes between 90 and 144 square meters, and those bigger than 144 square meters. Figure 5 plots the price indexes of these three categories in Shanghai and in Chongqing starting from January 2011. Apparently, for Shanghai the three curves are pretty much synchronized. But, for Chongqing, the price index for houses smaller than 90 square meters clearly divert from the other two to a higher level. Note that higher-end houses tend to be bigger; moreover, Chongqing houses smaller than 90 square meters are surely below the exemption size and thus free of property taxes while, for Shanghai, houses in each of the three categories may or may not be property-tax free, depending on the family size. The price patterns shown in Figure 5 point to a possibility that, in Chongqing, home purchasers turn to smaller houses instead so that the price of this category rises but that of others falls. This is consistent with our proposed explanation on the potential price spillover effect from higher-end to lower-end properties.

Nonetheless, the price patterns in Figure 5 do not serve as a test for the spillover effect. A complete test would involve the estimation of hypothetical prices of various property types in the absence of property taxes using the pre-treatment data. Unfortunately, the NBS price series of three housing categories begin with January 2011 only, and thus are not appropriate for this task.

6 Conclusion

This paper serves two purposes. It proposes an estimation approach similar in spirit to the conventional difference-in-difference approach, yet offers more flexibilities. This approach is most applicable to policy experiment and to non-stationary data. Thus, it can be an especially useful tool for evaluating macro policies.

Furthermore, it provides an important piece of suggestion for the intensive discussion on property taxes currently going on in China. That is, property taxation can be an effective policy tool to lower home prices, but should be carried out carefully. The case of Chongqing has presented a lesson, showing discriminative taxation based on property types can drive up home prices instead, possibly by causing a spillover effect from high-end to low-end properties. This should be taken into consideration in future property-tax design and implementation.

Future research should examine the time-series properties of the treatment effect of property taxes, when more post-treatment observations become available. It should also test for the spillover effect across various property types with more detailed data, to shed light on how property taxes influence the real estate market.

References

Bai, J, and Ng S. (2002). Determining the number of factors in approximate factor models. *Econometrica* 70, 191-221.

Cox, W. and H. Pavletich (2012). The 8th annual demographia international housing affordability survey. *Demographia*, Wendell Cox Consultancy, Illinois, U.S.A.

de Jong, R.M. and J. Davidson (2000) The functional central limit theorem and weak convergence to stochastic integrals I. *Econometric Theory* 16, 621-642.

Feldman, N. E.(2010). A reevaluation of Property Tax Capitalization: the Case of Michigan's Proposal A. Working Paper.

Hsiao, C., Ching, H.S. and S.K. Wand (2011). A panel data approach for program evaluaton: Measuring the benefit of political and economic integration of Hong Kong with mainland China. *Journal of Applied Econometricss*

Oates, W. E. (1969). The effects of property taxes and local public spending on property values: an empirical study of tax capitalization and the tiebout hypothesis. *Journal of Political Economy* 81(4), 1004-1008.

Palmon, D. and B. A. Smith (1998). New evidence on property tax capitalization. *Journal of Political Economy* 106(5), 1099-1111.

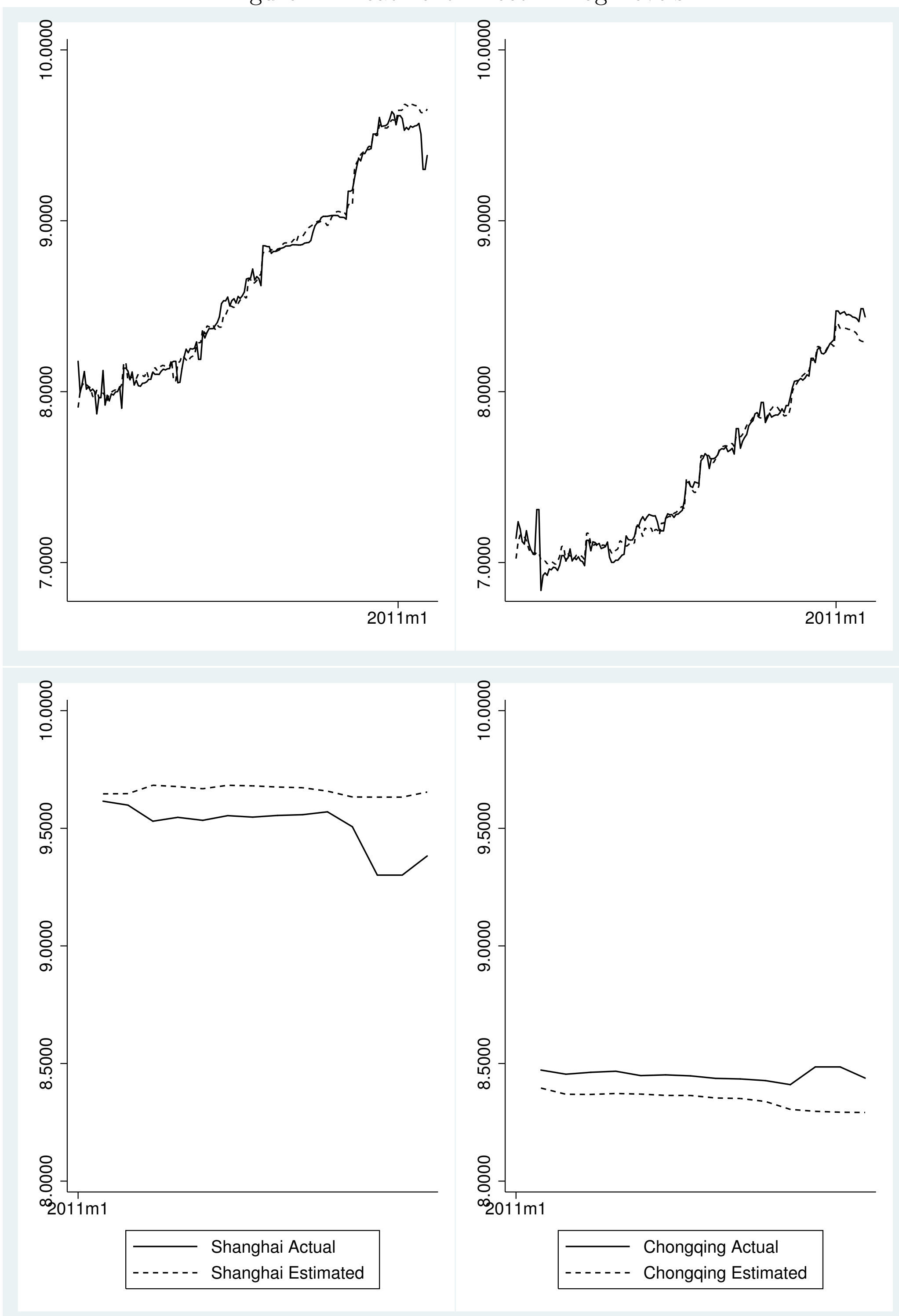
Poterba, J. M.(1984). Tax subsidies to owner-occupied housing: an asset-market approach. *Quarterly Journal of Economics* 99(4), 729-752.

Rosen, K T. (1982). The impact of proposition 13 on house prices in northern california: a test of the interjurisdictional capitalization hypothesis. *Journal of Political Economy* 90(1), 191-200.

Rosen, H. S. and D. J. Fullerton (1977) A note on local tax rates, public benefit levels, and property values. *Journal of Political Economy* 85(2), 433-440.

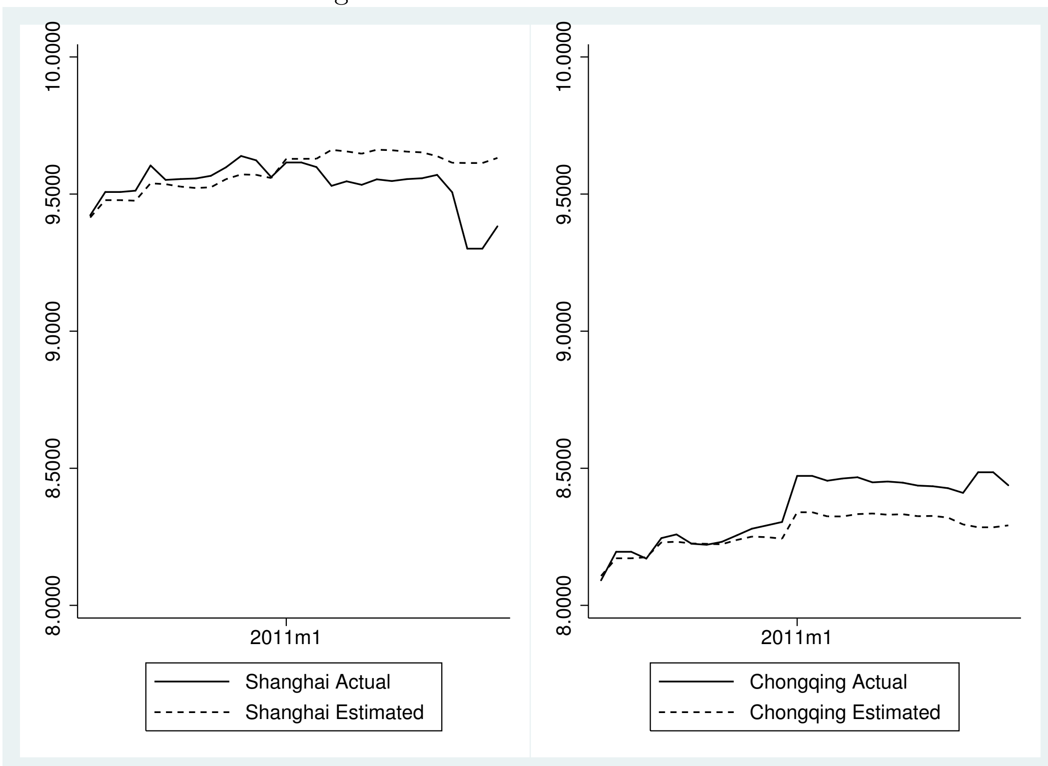
Xiao, Z. (2001). Likelihood-based inference in trending time series with a root near unity. *Econometric Theory* 17, 1113-1142.

Figure 1: Treatment Effect in Log Levels



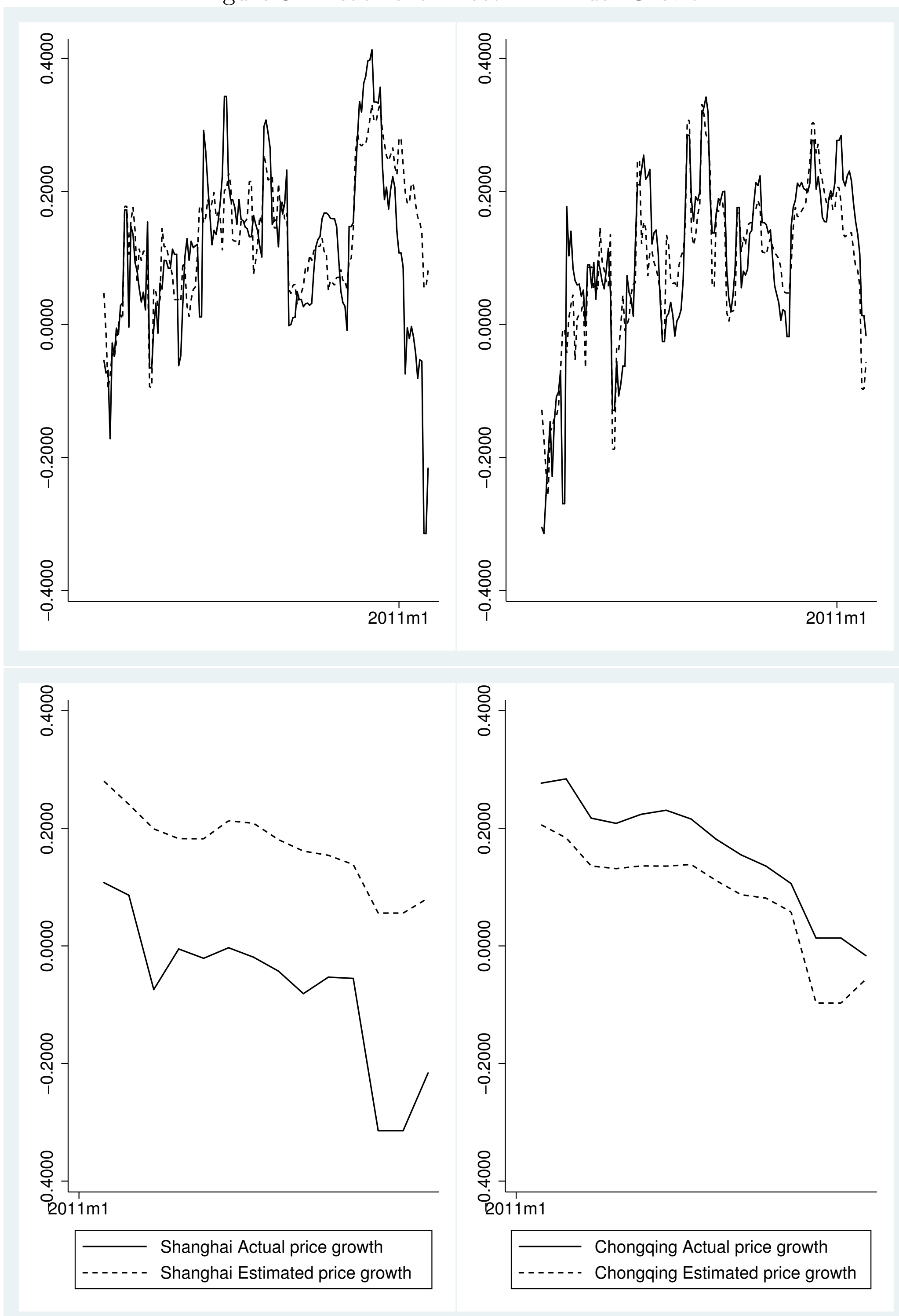
Note: data on actual price is from the National Development of Reform Committee (NDRC) of China; estimated price is based on data from March 1998 to January 2011. 2011m1 indicates January 2011 when the property-tax experiment is implemented. Prices are measured in log levels.

Figure 2: Robustness Check: the fit



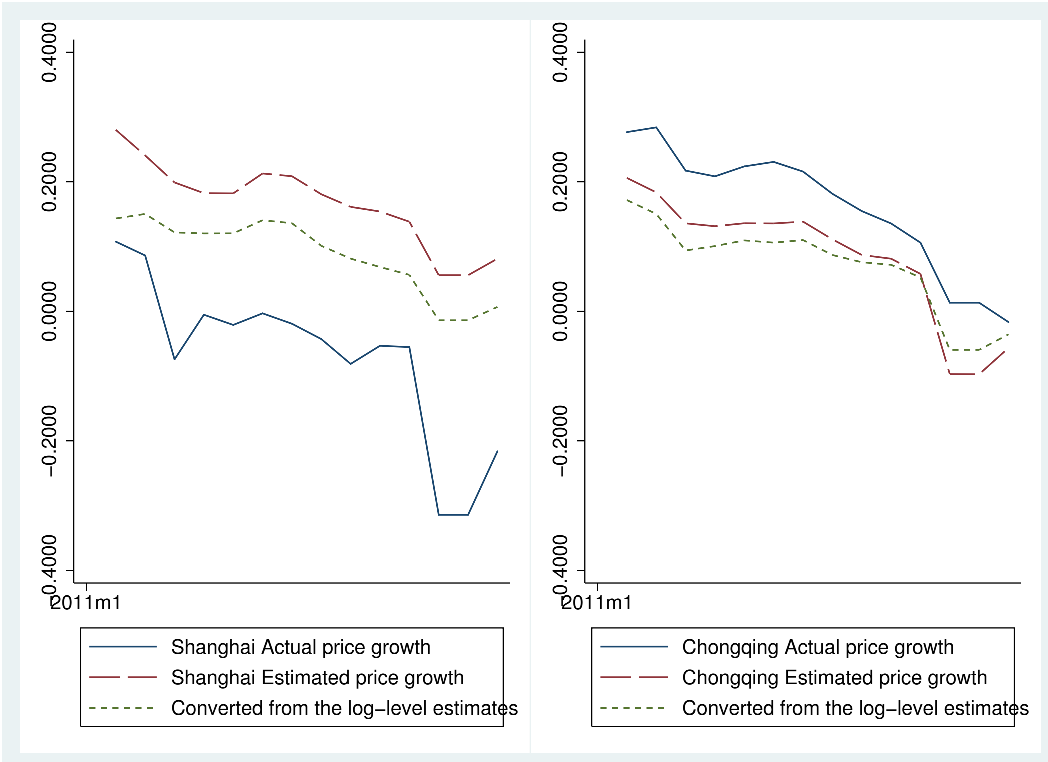
Note: Estimated price is based on data from March 1998 to November 2009. See Note for Figure 1 for other details.

Figure 3: Treatment Effect in Annual Growth



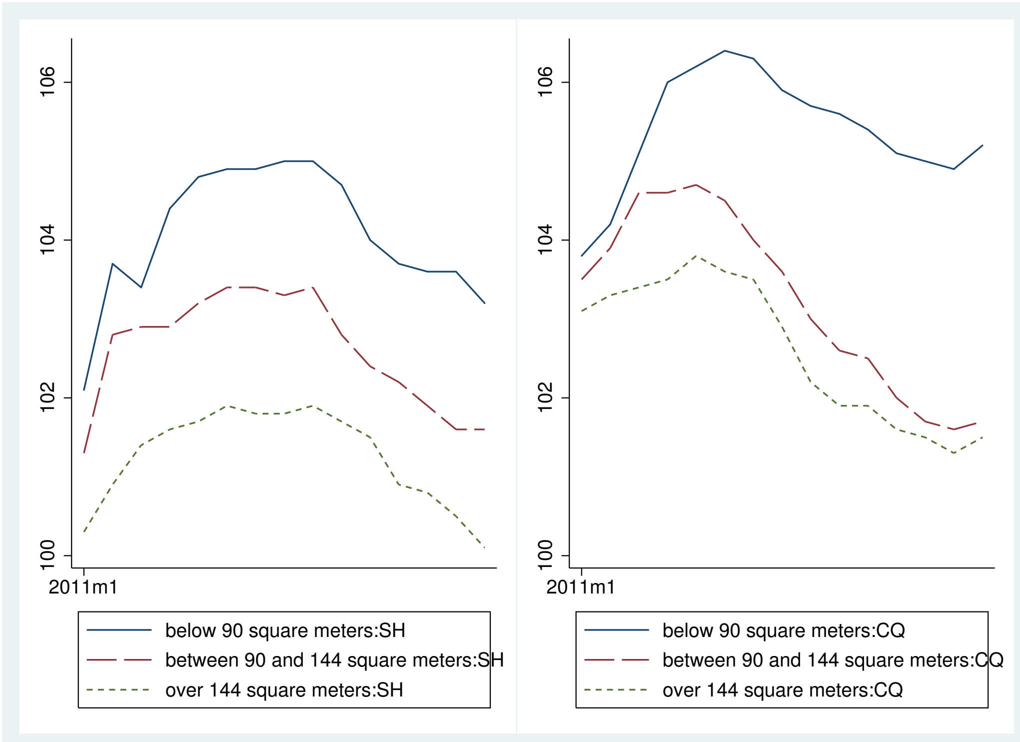
Note: Prices are measured in annual growth, as the differences in log levels compared with the corresponding month in the previous year. See Note for Figure 1 for other details.

Figure 4: Consistency Check: the log-level estimates and the annual growth estimates



Note: This figure checks the consistency between the log-level estimates and the annual-growth estimates. Estimated price growth is based on annual price growth data from March 1998 to January 2011. Growth is measured as the log difference between the present month and the corresponding month in the previous year. Short-dashed line indicates the annual growth converted from the estimated prices in log levels, namely, the difference between the estimated hypothetical prices in log level of the current month and those of the corresponding month in the previous year. 2011m1 indicates January 2011 when the property-tax experiment is implemented. See text for more details.

Figure 5: Price Trends of Three Home Types



Note: Price trends of three home types for Shanghai (SH) and Chongqing (CQ) starting from January 2011. Solid line indicates the fixed-indexed price level for homes smaller than 90 square meters; long-dashed line indicates that for homes between 90 and 144 square meters; short-dashed line indicates that of those bigger than 144 square meters. Data is from the National Bureau of Statistics (NBS) of China.

Table 1: Augmented Dickey-Fuller Tests for Unit Roots

Provinces/Cities	none	drift	trend	none	drift	trend
	log levels			log-first differences		
Beijing	0.8995	0.3228	0.5610	0.2400	0.0182*	0.7136
Tianjing	0.9744	0.5945	0.1042	0.0254*	0.0011*	0.0951*
Hebei	0.7239	0.1414	0.0175*	0.0000*	0.0000*	0.0000*
Liaoning	0.6557	0.1082	0.0001*	0.0012*	0.0000*	0.0110*
Shandong	0.9389	0.4222	0.0077*	0.0000*	0.0000*	0.0000*
Shanghai	.0.9098	0.3442	0.2005	0.2237	0.0165*	0.7653
Jiangsu	0.9863	0.7080	0.0203*	0.0017*	0.0001*	0.0191*
Zhejiang	0.9920	0.7944	0.0621*	0.0370*	0.0017*	0.2926
Fujian	0.9844	0.6852	0.5503	0.1160	0.0068*	0.2027
Guangdong	0.9768	0.6135	0.0853*	0.0194*	0.0008*	0.1522
Hainan	0.8771	0.2840	0.0010*	0.0000*	0.0000*	0.0002*
Guangxi	0.9341	0.4068	0.0521*	0.0039*	0.0001*	0.0326*
Jilin	0.0272	0.0013*	0.0000*	0.0000*	0.0000*	0.0000*
Heilongjiang	0.9401	0.4262	0.2473	0.0050*	0.0002*	0.0071*
Anhui	0.9729	0.5834	0.0780*	0.0018*	0.0001*	0.0126*
Jiangxi	0.9325	0.4019	0.1876	0.0007*	0.0000*	0.0005*
Hubei	0.9664	0.5419	0.0443*	0.0004*	0.0000*	0.0006*
Hunan	0.9500	0.4625	0.0104*	0.0262*	0.0011*	0.1264
Chongqing	0.9952	0.8632	0.1281	0.0058*	0.0002*	0.0221*
Sichuan	0.9909	0.7756	0.2595	0.0241*	0.0010*	0.0628*
Guizhou	0.9755	0.6030	0.3341	0.0042*	0.0002*	0.0091*
Yunan	0.8957	0.3156	0.0468*	0.0050*	0.0002*	0.0161*
Henan	0.9098	0.3442	0.0033*	0.0003*	0.0000*	0.0023*
Shanxi	0.9420	0.4326	0.0117*	0.0000*	0.0000*	0.0002*
ShanXi	0.2386	0.0181*	0.0000*	0.0000*	0.0000*	0.0001*
Neimenggu	0.3513	0.0327*	0.0118*	0.0000*	0.0000*	0.0001*
Xizang	0.3931	0.0410*	0.7991	0.0172*	0.0012*	0.0302*
Gansu	0.5014	0.0599*	0.0003*	0.0025*	0.0001*	0.0313*
Qinghai	0.8949	0.3144	0.0417*	0.0048*	0.0002*	0.0203*
Ningxia	0.9646	0.5313	0.2367	0.1966	0.0138*	0.4604
Xinjiang	0.9156	0.3573	0.0073*	0.0034*	0.0001*	0.0257*

Note: this table reports the Mackinnon approximated p values of the augmented Dickey-Fuller tests on home price of 31 provinces/cities in log levels and in log-first differences under various specifications. The log-first difference is the difference between log levels compared with the corresponding month in the previous year. *drift* indicates the test specification allows for a drift; *trend* implies the test includes a time trend; *none* means no drift or trend. All tests employ two lags, while the results are robust to various lag lengths. The sample size of each test is 166. Shanxi and ShanXi refer to two different provinces whose names share the same pronunciations. *indicates significance at the 10% level or above and therefore a rejection of the existence of a unit root.

con

Table 2: Weights of Control Provinces/Cities

Provinces/Cities	weights	SD	T	weights	SD	T
Panel A: Shanghai as the treatment city						
	$R^2 = 0.9907$			$R^2 = 0.9910$		
	F-stats=5299.69			F-stats=4145.84		
Jiangsu	0.7503***	0.0684	10.96	0.6453***	0.0909	7.10
Zhejiang	0.2908***	0.0668	4.35	0.2944***	0.0650	4.53
Heilongjiang	0.1976***	0.0551	3.59	0.1633***	0.0565	2.89
Sichuan	-0.2350***	0.0388	-6.06	-0.2157***	0.0385	-5.61
post-2008 dummy	0.1682***	0.0170	9.87	0.1753***	0.0175	10.00
time trend				0.0010	0.0007	1.50
Panel B: Chongqing as the treatment city						
	$R^2 = 0.9795$			$R^2 = 0.9837$		
	F-stats=4135.50			F-stats=3300.26		
Jiangsu	0.3973***	0.0688	5.77	0.7441***	0.0834	8.92
Zhejiang	0.1806***	0.0630	2.87	0.2044***	0.0601	3.40
Beijing	0.1232*	0.0656	1.88	0.1071*	0.0629	1.70
Sichuan	0.1342*	0.0810	1.66	0.1195*	0.0706	1.69
post-2008 dummy	0.0653***	0.0129	5.05	0.0623***	0.0132	4.73
time trend				-0.0035***	0.0007	-4.95

Note: This table reports the estimated γ' together with standard errors and t-statistics on $P_{1t}^0 = \gamma_1 + \gamma' \tilde{P}_t + \epsilon_{1t}$ with an optional time trend using home price before the policy experiment. P_{1t} is the average home price in Shanghai in Panel A and that in Chongqing in Panel B. Column 1 lists the provinces/cities whose average home price are included in $\tilde{P}_t$. Columns 2-4 report the estimation results without a time trend; columns 5-7 report those with a time trend. The sample size is 155. All estimations are conducted in log levels. ***indicates significance at the 1% level; *indicates significance at the 10% level.

Table 3: Treatment Effect of Property Tax

	Shanghai			Chongqing		
	Actual	Hypothetical	Treatment	Actual	Hypothetical	Treatment
			Panel A: with no time trend			
2011:m2	9.6151	9.6464	-0.0312	8.4721	8.3608	0.1114
2011:m3	9.5985	9.6472	-0.0488	8.4546	8.3445	0.1101
2011:m4	9.5300	9.6830	-0.1530	8.4626	8.3435	0.1191
2011:m5	9.5465	9.6771	-0.1307	8.4672	8.3523	0.1148
2011:m6	9.5337	9.6682	-0.1344	8.4486	8.3543	0.0943
2011:m7	9.5537	9.6827	-0.1290	8.4517	8.3493	0.1024
2011:m8	9.5474	9.6804	-0.1330	8.4474	8.3515	0.0959
2011:m9	9.5546	9.6753	-0.1207	8.4369	8.3442	0.0927
2011:m10	9.5577	9.6725	-0.1148	8.4344	8.3453	0.0890
2011:m11	9.5699	9.6579	-0.0880	8.4273	8.3392	0.0882
2011:m12	9.5066	9.6331	-0.1265	8.4101	8.3133	0.0968
2012:m1	9.3009	9.6325	-0.3316	8.4854	8.3013	0.1841
2012:m2	9.3009	9.6325	-0.3316	8.4854	8.3013	0.1841
2012:m3	9.3824	9.6540	-0.2717	8.4379	8.3090	0.1289
Average	9.5070	9.6602	-0.1532	8.4515	8.3364	0.1151
			Panel B: with a time trend			
2011:m2	9.6151	9.6415	-0.0264	8.4721	8.3956	0.0765
2011:m3	9.5985	9.6439	-0.0455	8.4546	8.3693	0.0853
2011:m4	9.5300	9.6761	-0.1461	8.4626	8.3686	0.0940
2011:m5	9.5465	9.6713	-0.1248	8.4672	8.3724	0.0947
2011:m6	9.5337	9.6641	-0.1304	8.4486	8.3699	0.0787
2011:m7	9.5537	9.6787	-0.1250	8.4517	8.3645	0.0872
2011:m8	9.5474	9.6774	-0.1299	8.4474	8.3641	0.0833
2011:m9	9.5546	9.6732	-0.1186	8.4369	8.3535	0.0834
2011:m10	9.5577	9.6713	-0.1136	8.4344	8.3513	0.0831
2011:m11	9.5699	9.6587	-0.0888	8.4273	8.3386	0.0887
2011:m12	9.5066	9.6367	-0.1301	8.4101	8.3049	0.1052
2012:m1	9.3009	9.6358	-0.3349	8.4854	8.2966	0.1888
2012:m2	9.3009	9.6366	-0.3357	8.4854	8.2929	0.1925
2012:m3	9.3824	9.6563	-0.2739	8.4379	8.2917	0.1462
Average	9.5070	9.6587	-0.1517	8.4515	8.3452	0.1063

Note: This table reports the estimated treatment effect after the implementation of property-tax experiment in January 2011 (2011:m1), as the difference between the actual price and the hypothetical price approximated using weights listed in Table 2. Panel A lists those with weights estimated with no time trend; Panel B lists those with weights estimated with a time trend.

con

Table 4: Weights from Estimation in Growth

Provinces/Cities	weights	SD	T	weights	SD	T
			Panel A: Shanghai as the treatment city			
	$R^2 = 0.8634$			$R^2 = 0.8332$		
	F-stats=158.44			F-stats=201.01		
Jiangsu	0.4594***	0.0882	5.21	0.4486***	0.0902	4.97
Zhejiang	0.4509***	0.0646	6.98	0.4358***	0.0700	6.23
Anhui	0.3385***	0.0706	4.79	0.3207***	0.0748	4.29
Sichuan	-0.3145***	0.0466	-6.76	-0.3111***	0.0464	-6.70
post-2008 dummy	0.1029***	0.0179	5.76	0.1033***	0.0181	5.72
time trend				0.0000	0.0000	0.55
			Panel B: Chongqing as the treatment city			
	$R^2 = 0.8001$			$R^2 = 0.8332$		
	F-stats=224.60			F-stats=201.01		
Jiangsu	0.1951	0.1876	1.04	0.3136*	0.1609	1.95
Zhejiang	0.2727***	0.0626	4.36	0.4129***	0.0669	6.17
Guangxi	0.2187**	0.1017	2.15	0.3936***	0.1094	3.60
Anhai	0.2851***	0.0884	3.22	0.2830***	0.0810	3.49
Jiangxi	-0.2297***	0.0663	-3.46	-0.1537**	0.0678	-2.27
Sichuan	0.1583**	0.0734	2.16	0.1172*	0.0667	1.76
post-2008 dummy	0.0611***	0.0121	5.06	0.0551***	0.0123	4.49
Time trend				-0.0001***	0.0000	-5.47

Note: This table reports the estimated γ' together with standard errors and t-statistics on $P^0_{1t} = \gamma_1 + \gamma' \tilde{P}_t + \epsilon_{1t}$ with an optional time trend measuring home price in log first differences (growth rates). Growth is measured as annual growth, as the change in log levels compared with the corresponding month in the previous year. The sample size is 143. All estimations are conducted in log levels. ***indicates significance at the 1% level; *indicates significance at the 10% level.

Table 5: Treatment Effect in Growth Rate

	Shanghai			Chongqing		
	Actual	Hypothetical	Treatment	Actual	Hypothetical	Treatment
			Panel A: with no time trend			
2011m2	0.1076	0.2802	-0.1726	0.2767	0.2007	0.0760
2011m3	0.0862	0.2407	-0.1545	0.2840	0.1881	0.0959
2011m4	-0.0742	0.1990	-0.2732	0.2172	0.1621	0.0551
2011m5	-0.0053	0.1824	-0.1876	0.2084	0.1613	0.0471
2011m6	-0.0211	0.1821	-0.2031	0.2238	0.1633	0.0605
2011m7	-0.0031	0.2128	-0.2159	0.2307	0.1524	0.0782
2011m8	-0.0192	0.2086	-0.2278	0.2158	0.1559	0.0599
2011m9	-0.0429	0.1807	-0.2236	0.1815	0.1369	0.0446
2011m10	-0.0812	0.1612	-0.2424	0.1548	0.1203	0.0345
2011m11	-0.0532	0.1541	-0.2073	0.1356	0.1198	0.0158
2011m12	-0.0553	0.1384	-0.1936	0.1060	0.1003	0.0057
2012m1	-0.3143	0.0558	-0.3700	0.0133	-0.0075	0.0208
2012m2	-0.3143	0.0558	-0.3700	0.0133	-0.0075	0.0208
2012m3	-0.2161	0.0812	-0.2973	-0.0167	0.0157	-0.0324
average	-0.0719	0.1666	-0.2385	0.1603	0.1187	0.0416
			Panel B: with a time trend			
2011m2	0.1076	0.2803	-0.1727	0.2767	0.2059	0.0708
2011m3	0.0862	0.2420	-0.1559	0.2840	0.1836	0.1004
2011m4	-0.0742	0.2021	-0.2763	0.2172	0.1358	0.0814
2011m5	-0.0053	0.1859	-0.1911	0.2084	0.1314	0.0770
2011m6	-0.0211	0.1855	-0.2065	0.2238	0.1359	0.0878
2011m7	-0.0031	0.2153	-0.2185	0.2307	0.1356	0.0950
2011m8	-0.0192	0.2111	-0.2303	0.2158	0.1385	0.0773
2011m9	-0.0429	0.1840	-0.2269	0.1815	0.1111	0.0703
2011m10	-0.0812	0.1653	-0.2465	0.1548	0.0870	0.0678
2011m11	-0.0532	0.1584	-0.2115	0.1356	0.0812	0.0544
2011m12	-0.0553	0.1433	-0.1986	0.1060	0.0578	0.0482
2012m1	-0.3143	0.0647	-0.3789	0.0133	-0.0970	0.1103
2012m2	-0.3143	0.0647	-0.3789	0.0133	-0.0972	0.1104
2012m3	-0.2161	0.0886	-0.3047	-0.0167	-0.0565	0.0398
average	-0.0719	0.1708	-0.2427	0.1603	0.0824	0.0779

Note: This table reports the estimated treatment effect measuring home price in growth rates. Growth is measured as annual growth, as the change in log levels compared with the corresponding month in the previous year. Panel A lists those with weights estimated with no time trend; Panel B lists those with weights estimated with a time trend.

Appendix A: Proofs of Propositions 2.1 to 2.3

Proof of Proposition 2.1.

The factor model is given by

$$(P_t)_{N\times 1} = (\alpha)_{N\times 1} + (B)_{N\times K}(f_t)_{K\times 1} + (u_t)_{N\times 1}, \tag{17}$$

where P_t, f_t are I(1) variables and u_t is a zero mean $I(0)$ variable. Assumption 1 implies that the K factors f_t are not cointegrated among themselves. Below we show that there exists cointegration vector $a = (1, -\gamma')'$ such that $a'P_t$ is I(0), i.e., P_t is cointegrated.

Premultiply P_t by a', we obtain

$$a'P_t = P_{1t} - \gamma'\tilde{P}_t = a'\alpha + a'Bf_t + a'u_t. \tag{18}$$

If we want to minimize $T_1^{-1}\sum_{t=1}^{T_1}(P_{1t} - \gamma'\tilde{P}_t)^2 = T_1^{-1}\sum_{t=1}^{T_1}(a'\alpha + a'Bf_t + a'u_t)^2$. It is obvious in order to minimize the above objective function, if the null space of B is non-empty, we will choose $a \in \mathcal{B}$, i.e., $a'B = 0$. This is because f_t is not cointegrated, if $a'B \neq 0$, then the $a'Bf_t$ term will make the objective function explode as $T_1 \to \infty$. Below we show that there exists $a = (1, \gamma')' \in \mathcal{R}^N$ such that $a'B = 0$. Without loss of generality we assume that the $K \times K$ square matrix B_K (see assumption 2) consists of the last K rows of $\tilde{B}$. Let γ_K denote the last K element of γ, and put $\gamma_j = 0$ for $j = 2, ..., N-1-K$. Then we have

$$a'B = (1, -\gamma')\begin{pmatrix} b_1' \\ \tilde{B} \end{pmatrix} = (b_1' - \gamma'\tilde{B}) = b_1' - \gamma_K' B_K \overset{set}{=} 0. \tag{19}$$

Since B_K is invertible, solving for γ_k from (19) leads to the unique solution of γ_K given by

$$\gamma_K = (B_K')^{-1} b_1.$$

Re-arranging terms in (18) and using (19) we get

$$P_{1t} = \gamma_1 + \gamma'\tilde{P}_t + \epsilon_{1t}, \tag{20}$$

where $\gamma_1 = a'\alpha$, $\epsilon_{1t} = a'u_t$. Because ϵ_{1t} is a stationary I(0) variable, (20) implies that P_{1t} and $\tilde{P}_t$ are cointegrated. OLS estimation based on (20) consistently estimate γ_1 and γ. This completes the proof of Proposition 2.1.

Proof of Proposition 2.2:

When f_t and P_t are unit root processes with drifts, we can re-write the regression model

$$P_{1t} = \gamma_1 + \gamma'\tilde{P}_t + \epsilon_{1t}$$

as

$$P_{1t} = \gamma_1 + \delta t + \gamma'\tilde{P}_t^* + \epsilon_{1t}, \qquad t = 1, ..., T_1, \tag{21}$$

where δ is a constant, $\tilde{P}_t^*$ is the de-tended process from $\tilde{P}_t$. Then it is well established (cite a few references here**???**) that $\hat{\gamma}_1 - \gamma_1 = O_p(T_1^{-1/2})$, $\hat{\delta} - \delta = O_p(T_1^{-3/2})$ and $\hat{\gamma} - \gamma = O_p(T_1^{-1})$, where $\hat{\gamma}_1$, $\hat{\delta}$ and $\hat{\gamma}$ are the OLS estimators of γ_1, δ and γ, respectively. Then we have

$$\begin{aligned} \hat{\Delta}_{1t} &= P_{1t}^1 - \hat{P}_{1t}^0 \\ &= P_{1t}^1 - P_{1t}^0 + P_{1t}^0 - \hat{P}_{1t}^0 \\ &= \Delta_{1t} + (\gamma_1 - \hat{\gamma}_1) + (\delta - \hat{\delta})t + \tilde{P}_t^{*'}(\gamma - \hat{\gamma}) + \epsilon_{1t} \\ &= \Delta_{1t} + \epsilon_{1t} + O_p(T_1^{-1/2}). \end{aligned}$$

We explain the above result in more details below. First, $\hat{\gamma} - \gamma = O_p(T_1^{-1/2})$. Next, since the de-tended variable $\tilde{P}_t^*$ is a unit root process without drifts, we know that $\hat{\gamma} - \gamma = O_p(T_1^{-1})$. Finally, $\hat{\delta} - \delta = O_p(T_1^{-3/2})$.

Hence, $\tilde{P}_t^{*'}(\gamma - \hat{\gamma}) = O_p(T^{1/2}T_1^{-1}) = O_p(T_1^{-1/2})$ since $T/T_1 = O(1)$ and $t(\hat{\delta} - \delta) = O_p(TT_1^{-3/2}) = O_p(T_1^{-1/2})$ by assumption 4. Also, we assume that there are no multicollinearity among different components of of P_{jt}^* for $j = 2, ..., N$, because we can always remove some of the regressors so that the remaining regressors do not suffer (severe) multicollinearity.

Proof of Proposition 2.3:

We will only prove the case that $\tilde{P}_t$ is a unit root process (without drift) since the case with drifts can be similarly proved.

$$\begin{aligned}
\hat{\Delta}_1 &= \frac{1}{T_2} \sum_{t=T_1+1}^{T} [P_{1t}^1 - \hat{P}_{1t}^0] \\
&= \frac{1}{T_2} \sum_{t=T_1+1}^{T} [P_{1t}^1 - P_{1t}^0 + P_{1t}^0 - \hat{P}_{1t}^0] \\
&= \frac{1}{T_2} \sum_{t=T_1+1}^{T} \Delta_{1t} + (\gamma_1 - \hat{\gamma}_1) + \frac{1}{T_2} \sum_{t=T_1+1}^{T} \tilde{P}_t'(\gamma - \hat{\gamma}) + \frac{1}{T_2} \sum_{t=T_1+1}^{T} \epsilon_{1t} \\
&= \Delta_1 + O_p(T_2^{-1/2}) + O_p(T_1^{-1/2}) + \left[\int_0^1 W(r)dr + o_p(1) \right] O_p(T_2^{1/2} T_1^{-1}) + O_p(T_2^{-1/2}) \\
&= \Delta_1 + O_p(\max\{T_1^{-1/2}, T_2^{-1/2}\})
\end{aligned}$$

because $T_2/T_1^2 = O(T_1^{-1})$, where we have also used $T_2^{-3/2} \sum_{t=T_1+1}^{T} \tilde{P}_t = T_2^{-1} \sum_{j=1}^{T_2} [\tilde{P}_j / T_2^{1/2}] \xrightarrow{d} \int_0^1 W(r)dr$ by assumption 3, and $\hat{\gamma} - \gamma = O_p(T_1^{-1})$ because $\tilde{P}_t$ is a driftless unit root process, and that P_{1t} and $\tilde{P}_t$ are cointegrated.